## APPENDIX D: SAMPLE SUDAAN PROCEDURE STATEMENTS

There are a number of releases of the SUDAAN software, running on several different platforms. Although the same procedure statements are used, there can be enhancements or subtle differences from one release to the next, particularly in reading and writing external data files. The statements displayed in the examples in this appendix are tailored for SUDAAN Release 8.01, SAS-Callable for Windows. The user should take this into consideration when using these examples or parts of these examples verbatim.

The example procedures represent relatively simple, straightforward applications. The options (various parameters, test statistics, etc.) in the sample programs may not be suitable for all your needs. Likewise, particular types of analyses may require options that are not displayed in the sample program statements. Our intention is not to suggest analytical approaches but to provide the key parameters that capture the relevant characteristics of the sample design. These parameters are found in the SUDAAN *design, weight,* and *nest* statements. In addition, the examples are limited to simple descriptive procedures for producing means or percentages. The same sample design parameters used for descriptive procedures are used for more complex estimation procedures such as regression or logit.
Preprocessing or recoding may be required for some variables because of missing or nonpositive data. Missing data are assigned SAS missing values (ex.: ".L – Legitimate missing," see Chapter 5.C - Value Coding Conventions). Classification (SUBGROUP) variables with zero or negative values will be treated by SUDAAN as missing and dropped from the procedure. This does not hold true for analysis variables (VAR) where zero or negative values are valid.

In using SUDAAN, the full sample should be processed even when analyses are limited to subgroups or subpopulations. This is to ensure the correct computation of the sampling variance. The SUDAAN statement SUBPOPN should be used to identify the specific analytic subpopulation of interest. The sampling variance estimates SUDAAN computes may be wrong if the file is reduced to a specific subpopulation prior to running the procedure.

**Example 1 – Variance Estimation Using Taylor Series Linearization**

This example shows the number of persons older than 16 by employment status and by whether they ever received SSI benefits. Standard errors of the percentages and unweighted and weighted population counts are included in the output. Note that C_SSI_EVER, a "0/1" dichotomous variable, has been recoded to "1/2" to conform to SUDAAN conventions for SUBGROUP variables. A SAS Format has also been created to applied value labels to the recoded variable. The SUBPOPN statement is used to identify persons older than 16 years of age and the file must be sorted by the NEST variables PseudoStrata and PseudoPSU.

```
PROC FORMAT;
  VALUE C_EVER
     1 = '1=No'
     2 = '2=Yes';
RUN;

PROC CROSSTAB DATA=final FILETYPE=SAS DESIGN=wr;
    SUBPOPN P_AGE>16;
    WEIGHT WgtFinal;
    NEST PseudoStrata PseudoPSU;
    SUBGROUP C_SSI_EVER C_SM_EMPLOYMENT;
    RECODE C_SSI_EVER=(0 1);
    RFORMAT C_SSI_EVER C_EVER. ;
    LEVELS 2 3;
    TABLE C_SSI_EVER C_SM_EMPLOYMENT;
    PRINT nsum wsum rowper serow;
    RUN;
```

**Example 2 – Variance Estimation Using Balanced Repeated Replication (BRR)**

Example 2 is a repeat of Example 1, but using Balanced Repeated Replication (BRR) rather than the Taylor Series Linearization procedure when calculating variances.

```
PROC FORMAT;
  VALUE C_EVER
     1 = 'No'
     2 = 'Yes';
RUN;

PROC CROSSTAB DATA=final FILETYPE=SAS DESIGN=BRR;
    SUBPOPN P_AGE>16;
    WEIGHT WgtFinal;
    REPWGT brr_wt1-brr_wt72/adjfay=4;
    SUBGROUP C_SSI_EVER C_SM_EMPLOYMENT;
   RECODE C_SSI_EVER=(0 1);
                RFORMAT C_SSI_EVER C_EVER. ;
    LEVELS 2 3;
    TABLE C_SSI_EVER C_SM_EMPLOYMENT;
    PRINT nsum wsum rowper serow;
    RUN;
```